

We thank the two reviewers for their careful reading of the manuscript and their comments. I have tried to respond in a way that both addresses the comments and improves the paper. In particular, the reviews suggest emphasizing what is innovative, important and/or significant, making the focus of the paper clearer, and eliminating redundant or unimportant parts. Since that has always been our goal, we are in fundamental agreement. That is described more completely below. There are also some smaller points where either the manuscript just needed a slight clarification or a mistake corrected, and a few where I disagree with the comments and have attempted to explain why.

Anonymous referee #1

The paper by Hintsa et al. attempts to present the NOAA UCATS instrument and discuss its measurement quality. The instrument was deployed on a series of missions over the past 15 years and has evolved considerably during this period. This latter fact presents the main problem I have with the paper. It does not describe one state of the instrument to a degree where this could be seen as a documentation or possibly as a reference for other instrument developers. Rather the paper describes in partly qualitative form the evolution of the instrument over time, partly in a narrative manner. As this evolution spans 15 years, it is virtually impossible to document everything with the desired accuracy. My suggestion would, however, be to add more detail to the description of the most recent set-up. In particular I would like to see more quantitative description of some of the GC parameters (column lengths, flow rates, temperatures) and also chromatograms of the GC channels. In particular I missed information on detector nonlinearity for the ECD channels and also on in-flight calibration frequency and in-flight precision vs. laboratory precision. If the authors estimate all this to be too much detail, I would also be happy to see this as supplementary material. I also suggest that table A1 should be moved to the main part of the publication, whereas the remaining figures in the Appendix could also be moved to a supplement. Next to these more general comment I have a couple of more specific comments given below. Overall the paper fits well into the scope of AMT and I believe it should be published once these points have been addressed.

The reviewer makes a good point here. I tried to emphasize the most recent set-up, but that obviously did not come through well enough, or was masked by other details. We have rewritten sections of the manuscript to more fully describe the instrument as flown in the ATom mission, which is the most recent one for which we have data to present. We added the details suggested by the reviewer about GC parameters, including detector non-linearity (and an associated figure). Many of these details also apply to earlier missions, though there were often small adjustments to the chromatography over time. These changes are mostly in the main text, with some details of more specialized interest in the Appendix. There is still some description of the earlier versions, as well as future changes, but hopefully these are contained and not a distraction. (One exception is for ATTREX ozone data, where we hope to present the scientific analysis of TTL structure and composition in a future publication; a thorough discussion of data quality for this is included.)

p. 3. l. 14-18. I suggest to include some references to ECD detector doping here.

Yes; we added a reference on doping that specifically discusses the interaction of N₂O with CO₂, CH₄, and H₂. In UCATS, that also allows us to detect CO (with N₂O doping).

p. 3l. 35: is there a model Nr. for the Maycomm hygrometer?

The hygrometers were both custom built, and that is now noted in the text.

p.5. l. 34.: as such papers will be around for a longer time period, I suggest not to use term like "upcoming", as this will be outdated in a short time. Rather use planned for the year X. Similar issues are found at other places in the manuscript. (e.g. p.19., l. 29, p. 20., l. 10)).

Agreed. The original text was written as the extent and impact of the Covid-19 pandemic were just becoming fully realized, and it was unclear how this would affect our schedules. Dates are now inserted, and at present seem likely to be correct.

p.8. l. 48.: I find this "trend-correction" very problematic. While this may be o.k. for SF₆ (rather constant trend, very small chemical loss), it is not appropriate for N₂O, where due to chemical loss, such a "trend-correction" should be done using a percentage increase. In both cases, this is very crude and it should be clear that some differences may remain which are due to trends.

Yes, part of the reason to make this adjustment was to allow both data sets to be plotted on the same scale, with similar axis ranges. The basic processes of horizontal and vertical transport and photochemistry in the atmosphere that lead to the correlation are the same (though their magnitudes may have changed); other than that there is no profound scientific insight in this figure - it is just a way of easily viewing the two data sets. The increase in SF₆ emissions may be the larger source of the differences; N₂O has increased much more steadily (and slowly). And you are right that a percentage increase is more appropriate for N₂O. I went back and recalculated the scaled POLARIS data with a percentage change for N₂O and CH₄. The figures changed, and have been updated. Obviously I confused (or unintentionally misled) the reviewers here; the text has been changed to make it clear that that the POLARIS data are mainly for reference to show the similar precision as for ACATS (but faster data rate), as well as the similar (not intended to be identical) structure in the atmosphere.

p. 10. l. 22: and Fig 4: It seems to me that the QCLS data are systematically higher in the troposphere. Have you checked absolute calibration scales? Please also discuss possible ECD non-linearity as mentioned above. This might be able to explain the significant deviations at low N₂O values.

Yes, the QCLS data are systematically higher in the troposphere. An offset of approximately 1.2 ppb was systematically observed in the ATom and HIPPO campaigns between QCLS and PFP data. These offsets in QCLS data relative to PFP (the reference instrument on board in both missions) are attributed to the calibration procedure. In a new paper (now cited in this manuscript), Gonzalez et al. [2021] adjust QCLS data for the small differences in the troposphere compared to the PFP samples in ATom data; this reduces QCLS N₂O values by a little more than 1 ppb, and the figures with ATom data in this manuscript (Figures 5 and A3) have also been

changed. The HIPPO data have now also been corrected in a similar way. We account for ECD non-linearity in our calibrations, though we do rely on that non-linearity remaining constant in flight (and from flight to flight). We think that one possibility for the deviations at low N₂O is from the fact that the in-flight standards used with the QCLS in HIPPO typically spanned a range from ~300 ppb to 330-350 ppb; the focus was on obtaining the best data in the troposphere, and (low) stratospheric values of N₂O were obtained by extrapolation. The vast majority of the QCLS data in HIPPO is within the span of the high and low calibration gases, though the stratospheric data stand out in the figures because of their large range. We adjusted the text to try to make this clearer.

p. 12. l. 17 and Figure 5: The correlation between N₂O and SF₆ looks quite remarkable. Could the authors add some more discussion on this? Are these data from different hemispheres?

Yes, there is a lot of information in this plot. In the stratosphere (or stratospherically influenced air), older air has lower N₂O (from photochemical loss, and the slightly smaller N₂O mixing ratio in older air) and lower SF₆ (tropospheric SF₆ was lower at earlier times), leading to the positive correlation shown. (The oldest air is in the lower left, as in Figure 2.) Data from both hemispheres (~70 S to 80 N) are shown in the plot, and the data near 330 ppb N₂O (actually most of the data points) are from the troposphere. This part of the plot reflects the very strong latitudinal gradient in SF₆, with lower SF₆ in the southern hemisphere and higher SF₆ in the northern hemisphere. This leads to a much steeper apparent slope (N₂O also has a latitudinal gradient, but weaker, also with lower values in the SH). Transitions between the troposphere and the stratosphere lead to mixing lines between the two branches, with the lower SF₆ transitions (8.6-9.0 ppt) in the SH, the higher SF₆ (>9.5 ppt) in the NH, and the tropical transition near 9.1 ppt. Part of this has been added to the text.